Raymond Salzmann

12/13/2020

Homework 3

Problem 1:

$x_i | c_i \sim \text{Binomial}(20, \theta_{c_i}) \qquad c_i \overset{iid}{\sim} \text{Discrete}(\pi)$

$$\ln p(x|\pi,\theta) = \sum_c q(c) \ln \frac{p(x,c|\pi,\theta)}{q(c)} + \sum_c q(c) \ln \frac{q(c)}{p(c|x,\pi,\theta)}$$

E-Step: $q(c) = \ln p(c|x,\pi,\theta)$

$$p(c|x,\pi,\theta) \propto \prod_{i=1}^{N} p(x_i|c_i,\theta)\, p(c_i|\pi)$$

$$= \prod_{i=1}^{N} p(c_i|x_i,\pi,\theta)$$

$$= \prod_{i=1}^{N} q(c_i)$$

$$p(c_i=k|x_i,\pi,\theta) = \frac{p(x_i|c_i=k,\pi,\theta)\, p(c_i=k|\pi)}{\sum_{j=1}^{K} p(x_i|c_i=j,\pi,\theta)\, p(c_i=j|\pi)}$$

$$= \frac{\pi_k \,\text{Binomial}(x_i|\theta_k)}{\sum_{j=1}^{K} \pi_j \,\text{Binomial}(x_i|\theta_j)}$$

$$\mathcal{L}(\pi,\theta) = \sum_{i=1}^{N} \mathbb{E}_{q(c)}\left[\ln p(x_i,c_i|\pi,\theta)\right]$$

$$= \sum_{i=1}^{N} \sum_{j=1}^{K} \phi_i(j) \left[\ln p(x_i|c_i=j,\pi,\theta) + \ln p(c_i=j|\pi)\right]$$

$$= \sum_{i=1}^{N} \sum_{j=1}^{K} \phi_i(j) \left[x_i \ln(\theta_j) + (20 - x_i) \ln(1-\theta_j) + \ln(\pi_j) \right]$$

M-Step:

$$\nabla_{\theta_j} \mathcal{L}(\pi, \theta) = 0 \quad \text{where } \theta_j = \frac{\sum_{i=1}^{N} \phi_i(j) x_i}{20 \sum_{i=1}^{N} \phi_i(j)} = \frac{\sum_{i=1}^{N} \phi_i(j) x_i}{20 \, n_j}$$

And $n_j = \sum_{i=1}^{N} \phi_i(j)$

$$\nabla_{\pi} \mathcal{L}(\pi, \theta) = 0 \quad \text{where } \pi_j = \frac{n_j}{n} \quad \text{and} \quad n_j = \sum_{i=1}^{N} \phi_i(j)$$

$$b_t = \ln p(x \mid \pi, \theta) = \sum_{i=1}^{N} \sum_{j=1}^{K} \ln \pi_j + \ln \text{Binomial}(x_i \mid \theta_j)$$

Pseudo Code:

Given Data:
- $x_1, \ldots, x_n$; $x_i \in \mathbb{R}^d$
- K = number of clusters.

Output:
- Parameters π, θ
- Cluster Assignment Distributions

(1) Initialize the π, θ parameters in some way for the first iteration

(2) For every iteration, $t = 1, \ldots, T$:

(a) E-step: For $i = 1, \ldots, N$ and $j = 1, \ldots, K$:

$$\phi_i^{(t)}(j) = \frac{\pi_j \text{Binomial}(x_i \mid \theta_j)}{\sum_{k=1} \pi_k \text{Binomial}(x_i \mid \theta_k)}$$

(b) M-step:

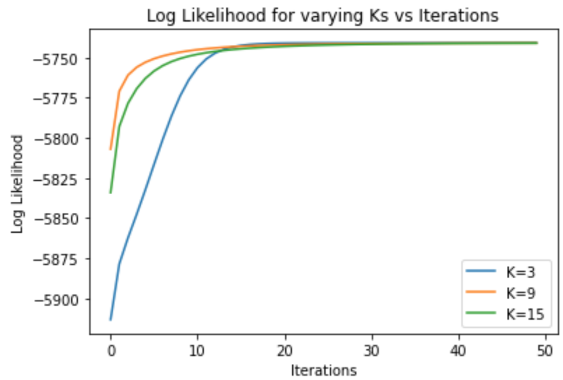
$$n_j^{(t)} = \sum_{i=1}^{N} \phi_i^{(t)}(j)$$

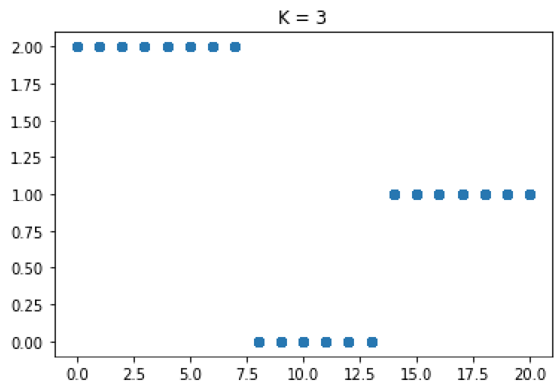
$$\theta_j = \frac{\sum_{i=1}^{N} \phi_i^{(t)}(j)\ x_i}{20\, n_j^{(t)}}$$

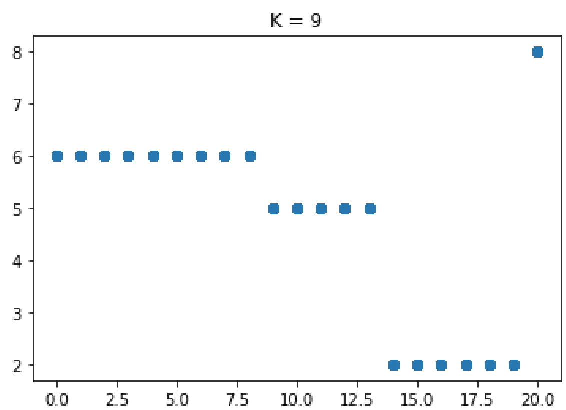
$$\pi_j = \frac{n_j^{(t)}}{n}$$

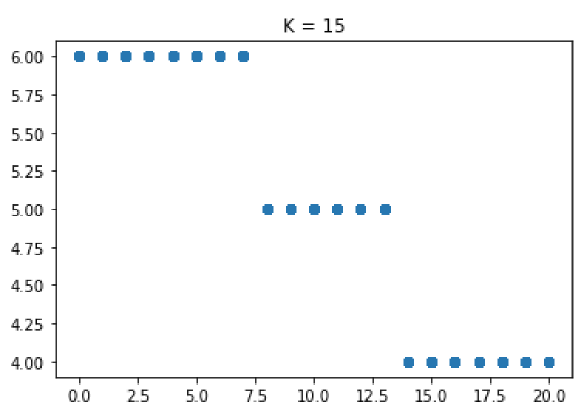
(3) Calculate log-likelihood:

$$b_t = \ln p(x \mid \pi, \theta)$$


Log Likelihood for varying Ks vs Iterations
Log Likelihood
−5750
−5775
−5800
−5825
−5850
−5875
−5900
Iterations
0
10
20
30
40
50
K=3
K=9
K=15


K = 3
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0


K = 9
8
7
6
5
4
3
2
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0


K = 15
6.00
5.75
5.50
5.25
5.00
4.75
4.50
4.25
4.00
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0

Problem 2:

$\pi \sim \text{Dirichlet}(\alpha)$, $\theta_k \overset{iid}{\sim} \text{Beta}(a,b)$

$$p(x,c,\pi,\theta) = \prod_{i=1}^{N} p(x_i, c_i, \pi, \theta)$$

$$= \left(\prod_{i=1}^{N} p(x_i \mid c_i, \pi, \theta)\, p(c_i \mid \pi)\right)\left(\prod_{j=1}^{K} p(\theta_j)\right) p(\pi)$$

q-distributions:

$$\mathbf{q(c_i=j)} \propto e^{\mathbb{E}_{-q}\left[\ln p(x_i, c_i = j \mid \pi, \theta_j)\right]}$$

$$\propto e^{\mathbb{E}_{-q}\left[\ln p(x \mid \theta_j)\right] + \mathbb{E}\left[\ln p(c_i = j \mid \pi)\right]}$$

$$\propto e^{x_i \mathbb{E}\left[\ln \theta_j\right] + (20 - x_i)\mathbb{E}\left[\ln(1-\theta_j)\right] + \mathbb{E}\left[\ln \pi_j\right]}$$

$$\propto e^{x_i\left[\psi(a_j) - \psi(a_j' + b_j')\right) + (20-x_i)\left(\psi(b_j') - \psi(a_j' + b_j')\right) + \psi(\alpha_j') - \psi\left(\sum_k \alpha_k\right)\right]}$$

$$\mathbf{q(\pi)} \propto e^{\sum_{i=1}^{N} \mathbb{E}\left[\ln p(c_i = j \mid \pi)\right] + \ln p(\pi)}$$

$$\propto e^{\sum_{i=1}^{N}\sum_{j=1}^{K} \phi_i(j) \ln \pi_j + \sum_{j=1}^{K} (\alpha - 1)\ln(\pi_j)}$$

$$\propto \prod_{j=1}^{K} \pi_j^{\sum_{i=1}^{N} \phi_i(j) + \alpha - 1}$$

$$= \text{Dirichlet}(\alpha')$$

$$\alpha_j' = n_j + \alpha$$

$$n_j = \sum_{i=1}^{N} \phi_i(j)$$

$$q(\theta_j) \propto e^{\sum_{i=1}^{N} \mathbb{E}[\ln p(x \mid c_i = j, \pi, \theta)] + \ln p(\theta_j)}$$

$$\propto e^{\left[\sum_{i=1}^{N} \phi_i(j) x_i + a - 1\right] \ln \theta_j + \left[\sum_{i=1}^{N} \phi_i(j)(20 - x_i) + b - 1\right] \ln(1-\theta_j)}$$

$$= \text{Beta}(a', b')$$

where

$$a' = \sum_{i=1}^{N} \phi_i(j) x_i + a$$

$$b' = \sum_{i=1}^{N} \phi_i(j)(20 - x_i) + b$$

Calculating the ELBO:

$$\mathcal{L} = \mathbb{E}[\ln p(x, c, \pi, \theta)] - \mathbb{E}[\ln q]$$

$$= \sum_{i=1}^{N} \sum_{j=1}^{K} \mathbb{E}\left[\mathbb{1}(c_i = j)(\ln p(x_i \mid \theta_j) + \ln \pi_j)\right] + \sum_{j=1}^{K} \mathbb{E}[\ln p(\theta_j)]$$

$$+ \mathbb{E}[\ln p(\pi)] - \sum_{i=1}^{N} \mathbb{E}[\ln q(c_i)] - \sum_{K=1}^{K} \mathbb{E}[\ln q(\theta_K)] - \mathbb{E}[\ln q(\pi)]$$

(i) $\mathbb{E}\left[\mathbb{1}(c_i = j)(\ln p(x_i \mid \theta_j) + \ln \pi_j)\right] = \phi_i(j)\, \mathbb{E}[\ln p(x_i \mid \theta_j) + \ln \pi_j]$

$$= \phi_i(j)\left(\mathbb{E}[\ln p(x_i \mid \theta_j)] + \mathbb{E}[\ln p(\pi_j)]\right)$$

$$= \phi_i(j)\Big(\ln(20C_{x_i}) + x_i\left[\psi(a_j') - \psi(a_j' + b_j')\right] + (20 - x_i)\big(\psi(b_j') -$$

$$\psi(a_j' + b_j')\big)\Big) + \psi(\alpha_j') - \psi\left(\sum_k \alpha_k'\right)$$

(2) $\mathbb{E}[\ln p\, \theta_j]$

$$= (a_j - 1)\left[\psi(a_j') - \psi(a_j' + b_j')\right] + (b_j - 1)\left[\psi(b_j') - \psi(a_j' + b_j')\right]$$

$$- \ln B(a_j, b_j)$$

(3) $\mathbb{E}[\ln p(\pi)] = \sum_i (\alpha_i - 1)\left[\psi(\alpha_i') - \psi\left(\sum_k \alpha_k'\right)\right] - \ln B(\alpha_j)$

(4) $\mathbb{E}[\ln q(c_i)] = \sum_{j=1}^{k} \phi_i(j) \ln \phi_i(j)$

(5) $\mathbb{E}[\ln q(\theta_j)] = (a_j' - 1)\left[\psi(a_j') - \psi(a_j' + b_j')\right] + (b_j' - 1)\left[\psi(b_j')\right.$

$$\left. - \psi(a_j' + b_j')\right] - \ln B(a_j', b_j')$$

(6) $\mathbb{E}[\ln q(\pi_j)] = \sum_i (\alpha_i' - 1)\left[\psi(a_i') - \psi\left(\sum_k \alpha_k'\right)\right] - \ln B(\alpha_j')$

For all above derivations (1) - (6):

$a' = a^{(t)}$, $b' = b^{(t)}$, $\alpha' = \alpha^{(t)}$

Pseudo-Code:

(1) Initialize $(\alpha_1^{(0)}, \ldots, \alpha_k^{(0)})$, $(a_j^{(0)}, b_j^{(0)})$ in some way to be used for the first iteration

(2) For all iterations $t = 1, \dots, T$

(a) Update $q(c_i)$ for $i = 1, \dots, N$

$$\phi_i^{(t)}(j) = \frac{e^{x_i \psi_1' + (20 - x_i)\psi_2' + \psi_3'}}{\sum_{k=1}^{K} e^{x_i \psi_1' + (20 - x_i)\psi_2' + \psi_3'}}$$

where $\psi_1' = \psi(a_j^{(t-1)}) - \psi(a_j^{(t-1)} + b_j^{(t-1)})$

$$\psi_2' = \psi(b_j^{(t-1)}) - \psi(a_j^{(t-1)} + b_j^{(t-1)})$$

$$\psi_3' = \psi(\alpha_j^{(t-1)}) - \psi(\sum_k \alpha_k^{(t-1)})$$

(b) Set $n_j^{(t)} = \sum_{i=1}^{N} \phi_i^{(t)}(j)$ for $j = 1, \dots, k$

(c) Update $q(\pi)$: $\alpha_j^{(t)} = \alpha + n_j^{(t)}$ for $j = 1, \dots, K$

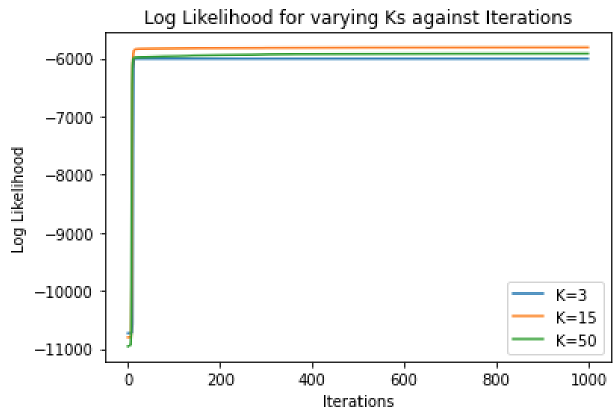
(d) Update $q(\theta_j)$:

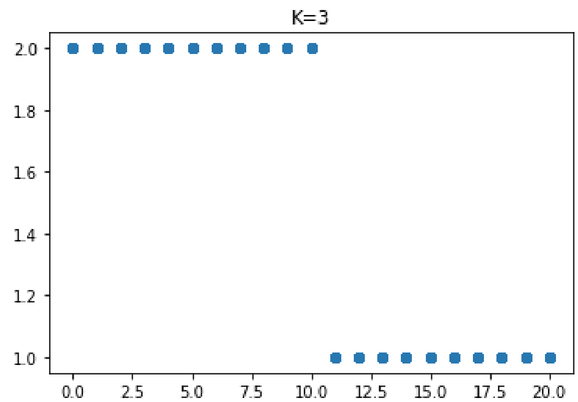
$$a^{(t)} = \sum_{i=1}^{N} \phi_i^{(t)}(j)\, x_i + a$$

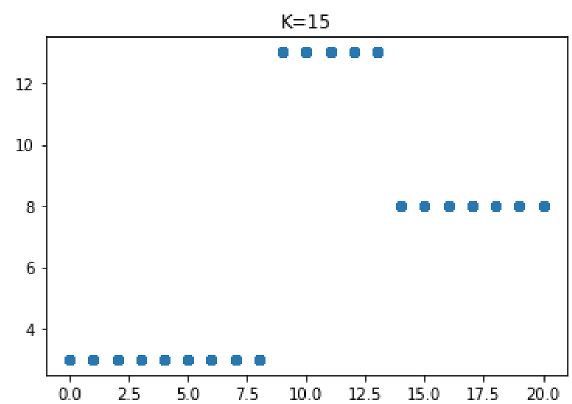
$$b^{(t)} = \sum_{i=1}^{N} \phi_i^{(t)}(j)\, (20 - x_i) + b$$

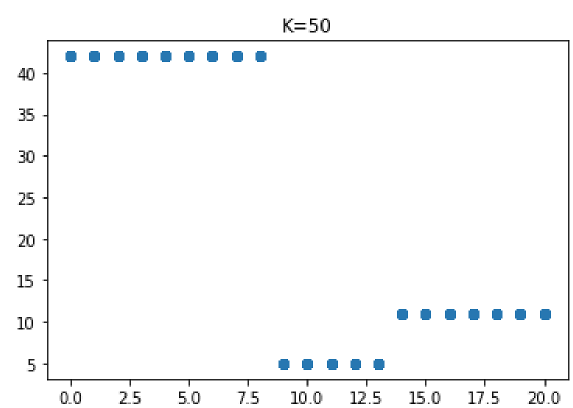
(e) Calculate the ELBO:

$$\mathcal{L} = \mathbb{E}[\ln p(x, c, \pi, \theta)] - \mathbb{E}[\ln q]$$


Log Likelihood for varying Ks against Iterations
Log Likelihood
Iterations
K=3
K=15
K=50


K=3


K=15


K=50

Problem 3:

$$p(\theta \mid \{x_i : c_i = j\}) \propto \left[\prod_{i=1}^{N} p(x_i \mid \theta_j)^{\mathbb{1}(c_i=j)}\right] p(\theta_j)$$

$$\propto \theta_j^{\sum_{i=1}^{N} \mathbb{1}(c_i=j) x_i} \cdot (1-\theta_j)^{\sum_{i=1}^{N} \mathbb{1}(c_i=j)(20-x_i)} \cdot \theta_j^{a-1} \cdot (1-\theta_j)^{b-1}$$

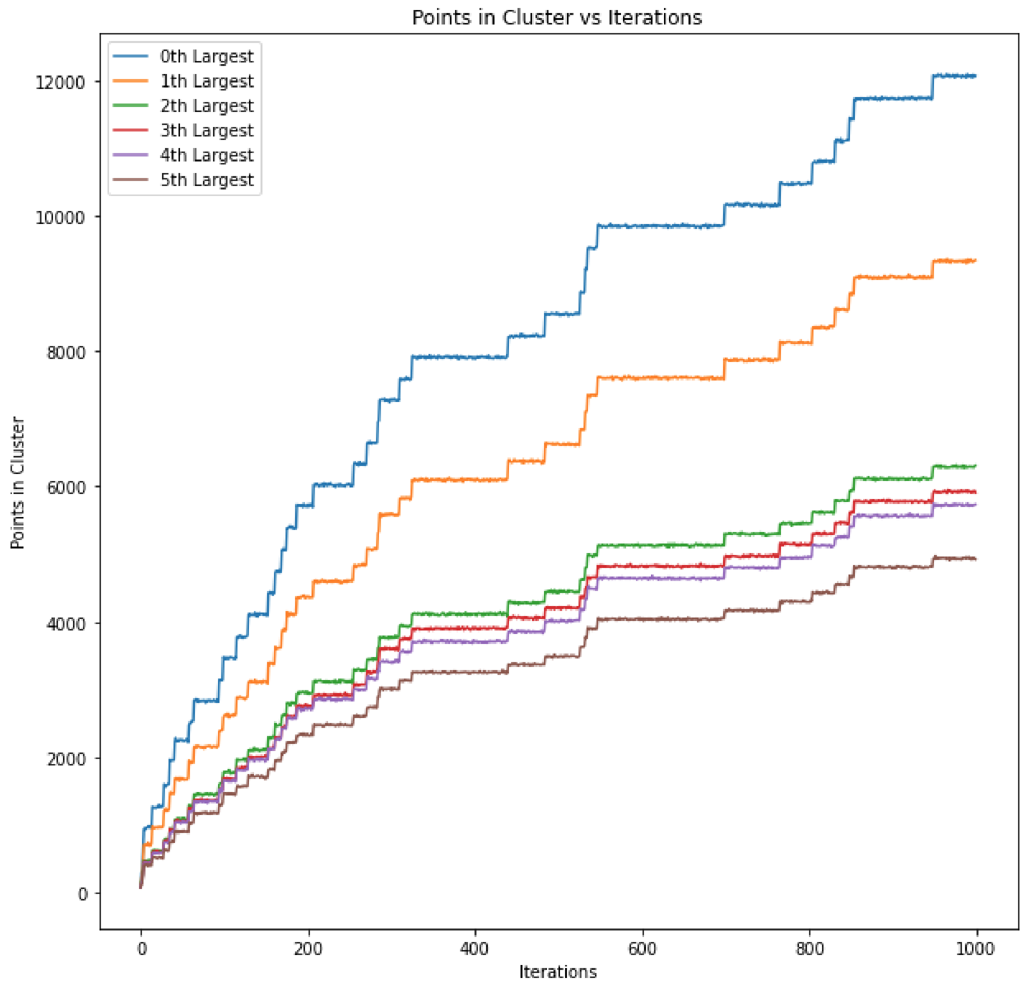
$$\propto \theta_j^{a+\sum_{i=1}^{N} \mathbb{1}(c_i=j) x_i} \cdot (1-\theta_j)^{b+\sum_{i=1}^{N} \mathbb{1}(c_i=j)(20-x_i)}$$

$$= \text{Beta}(a', b')$$

$$a' = a + \sum_{i=1}^{N} \mathbb{1}(c_i=j) x_i \quad , \quad b' = b + \sum_{i=1}^{N} \mathbb{1}(c_i=j)(20-x_i)$$

$$p(c_i = j \mid x_i, \theta, c_{-i}) \propto p(x_i \mid \theta_j) \cdot \frac{n_j(-i)}{\alpha + n - 1}$$

$$\propto \text{Binomial}(x_i \mid 20, \theta_j) \cdot \frac{n_j(-i)}{\alpha + n - 1}$$


Points in Cluster vs Iterations
0th Largest
1th Largest
2th Largest
3th Largest
4th Largest
5th Largest
Points in Cluster
0
2000
4000
6000
8000
10000
12000
0
200
400
600
800
1000
Iterations


Number of Clusters vs Iterations
Number of Clusters
30
31
32
33
34
35
0
200
400
600
800
1000
Iterations